DOCUMENT RESUME

ED 477 084 IR 021 812

AUTHOR Reeves, Thomas C.; Benson, Lisa; Elliott, Dean; Grant, Michael; Holschuh, Doug; Kim, Beaumie; Kim, Hyeonjin; Lauber, Erick; Loh, Sebastian

TITLE Usability and Instructional Design Heuristics for E-Learning Evaluation.

PUB DATE 2002-06-00

NOTE 8p.; In: ED-MEDIA 2002 World Conference on Educational Multimedia, Hypermedia & Telecommunications. Proceedings (14th, Denver, Colorado, June 24-29, 2002); see IR 021 687.

AVAILABLE FROM Association for the Advancement of Computing in Education (AACE), P.O. Box 3728, Norfolk, VA 23514. Tel: 757-623-7588; e-mail: info@aace.org; Web site: http://www.aace.org/DL/.

PUB TYPE Reports - Research (143) -- Speeches/Meeting Papers (150)

EDRS PRICE EDRS Price MF01/PC01 Plus Postage.

DESCRIPTORS Computer Assisted Instruction; *Computer Software Evaluation; Distance Education; Educational Technology; *Evaluation Criteria; *Evaluation Methods; *Heuristics; Higher Education; *Instructional Design

IDENTIFIERS *Usability

ABSTRACT

Heuristic evaluation is a methodology for investigating the usability of software originally developed by Nielsen (1993, 2000). Nielsen's protocol was modified and refined for evaluating e-learning programs by participants in a doctoral seminar held at the University of Georgia in 2001. The modifications primarily involved expanding Nielsen's original 10 heuristics (developed for software in general) to 15 heuristics (designed to be more closely focused on e-learning programs). The application of this evaluation protocol to a commercial e-learning program supported enhancements in the usability of the program. This paper describes the set of 15 e-learning heuristics as well as the protocol that guided the evaluation process. In addition, the results of the application of this heuristic evaluation protocol to a commercial e-learning training program are illustrated. (Author)

Reproductions supplied by EDRS are the best that can be made from the original document.

Usability and Instructional Design Heuristics for E-Learning Evaluation

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY

G.U. Marks

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

Thomas C. Reeves, Lisa Benson, Dean Elliott, Michael Grant, Doug Holschuh, Beaumie Kim, Hyeonjin Kim, Erick Lauber, and Sebastian Loh

Department of Instructional Technology, The University of Georgia
604 Aderhold Hall, Athens, GA 30602-7144 USA
Main Contact E-Mail: treeves@coe.uga.edu

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

✓ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

Abstract: Heuristic evaluation is a methodology for investigating the usability of software originally developed by Nielsen (1993, 2000). Nielsen's protocol was modified and refined for evaluating e-learning programs by participants in a doctoral seminar held at The University of Georgia in 2001. The modifications primarily involved expanding Nielsen's original ten heuristics (developed for software in general) to fifteen heuristics (designed to be more closely focused on e-learning programs). The application of this evaluation protocol to a commercial e-learning program supported enhancements in the usability of the program.

ED 477 084

Introduction

Heuristic evaluation is a methodology for investigating the usability of software originally developed by Jakob Nielsen (1993, 2000), a widely acknowledged usability expert. According to Nielsen (1994), heuristic evaluation "involves having a small set of evaluators examine the interface and judge its compliance with recognized usability principles (the "heuristics")." Nielsen's original protocol for heuristic evaluation can be found on the Web at: http://www.useit.com/papers/heuristic/.

For this study, Nielsen's protocol was modified and refined for evaluating e-learning programs by the participants in a doctoral seminar held at The University of Georgia between August and December 2001. The modifications primarily involved expanding Nielsen's original ten heuristics (developed for evaluating software in general) to fifteen heuristics (developed for evaluations of e-learning programs). This paper describes the set of fifteen e-learning heuristics as well as the protocol that guided the evaluation process. In addition, the results of the application of this heuristic evaluation protocol to a commercial e-learning training program are illustrated.

Method

Figure 1 presents the protocol we developed for the heuristic evaluation of e-learning programs. Figure 2 presents a set of fifteen usability and instructional design heuristics which should be viewed as a work in progress. Further refinements are likely to be made based upon the application of the heuristics to various e-learning products as well as the feedback we receive from others who use this tool for their own heuristic evaluations.

Seven instructional technology doctoral students and two faculty members (one from cognitive psychology and one from instructional technology) were involved in the development and initial application of these heuristics and the accompanying protocol. There were twenty heuristics in our initial set. These were generated through a critical analysis of Nielsen's (1994) heuristics as well as several small and whole group "brainstorming" sessions aimed at identifying additional heuristics for the usability evaluation of e-learning programs.

The initial twenty heuristics and the protocol were first applied to a commercial e-learning program called "GMP Basics" designed for the American Red Cross by LearnWright, an award-winning company located in Rockville, Maryland, USA (http://www.learnwright.com). LearnWright specializes in the development of e-learning programs for regulated industries such as pharmaceuticals. Training for these industries is mandated by the U.S. Food and Drug Administration (FDA). From the perspective of the FDA, the collection of blood and plasma products is a regulated activity. While not everyone at the Red Cross is involved in collecting, processing, or distributing blood or plasma products (e.g., some workers are involved exclusively in disaster relief), all workers in blood services must learn the

IR021812

BEST COPY AVAILABLE

principles encompassed in the FDA's regulations for Good Manufacturing Practices (GMP) The courseware we evaluated was presented on a CD-ROM, although it will eventually be accessible via a Web-based intranet.

Protocol for E-Learning Heuristic Evaluation

This instrument and protocol are intended for use by instructional designers and other experts engaged in heuristic evaluations of e-learning programs. The instrument itself lists fifteen heuristics for e-learning programs, some of which are based upon Jakob Nielsen's widely used protocol for heuristic evaluation of any type of software (http://useit.com/papers/heuristic/), and the rest of which are based upon factors related to instructional design. Although we have tried to be comprehensive, experts may decide to add new heuristics deemed relevant to the types of e-learning programs being evaluated or to the expert's specific expertise.

Steps:

1. An expert should review the heuristics and accompanying "Sample questions to ask yourself" in the instrument before reviewing an e-learning product. The expert should modify the instrument if needed, by adding, deleting, or changing heuristics.

2. It is important that the expert spend substantial time exploring the e-learning program before beginning the actual heuristic evaluation. Ideally, the expert will assume the role of typical learner who would use this e-learning program. Before beginning the review, the expert should be given (or try to discover) background information related to the e-learning program such as:

a. *Target audience and learner characteristics*: A thorough description of the intended audience and their learner characteristics (e.g., education level, motivation, incentive, and computer expertise) will enable the expert to judge the appropriateness of the user interface and other aspects of the program's usability in an informed manner.

b. *Instructional goals and objectives*: The expert should know as much as possible about the needs that the e-learning program is intended to address, ideally in terms of clear goals and objectives.

c. *Typical context for using this program*: Realistic scenarios for when, where, and how the e-learning program will be used should be known by the expert.

d. *Instructional design strategies used in the program*: If possible, a description of the design specifications used in developing the e-learning program should be provided to the expert so that the expert's judgment of the appropriateness of the instructional design strategies are informed with respect to the instructional designer's intentions.

e. *The status of the program's development and possibilities for change*: The expert should be informed as to where the program is in the development cycle (e.g., an early prototype, beta version, or older version under consideration for redesign).

3. After spending enough time to become familiar with the program, the expert should go through it from beginning to end to conduct the actual heuristic evaluation. (With lengthy programs, the expert may only review a representative sample of the program.)

4. The expert should make note of every usability problem found. For each problem, the expert should identify the heuristic it violates, and then give it a severity rating using the severity scale below. If the problem cannot be attributed to a violation of a specific heuristic, the expert should make a note of this. (If a number of problems are found that cannot be associated with specific heuristics, this may suggest the need for the development of new heuristics.)

Severity Scale

1. cosmetic problem only; need not be fixed unless extra time is available
2. minor usability problem; fixing this should be given low priority
3. major usability problem; important to fix; so should be given a high priority
4. usability catastrophe; imperative to fix before this product is released

5. After all the usability problems are found, the expert should go back though them and give each one an extensiveness rating using the extensiveness scale below.

Extensiveness Scale

1. this is a single case
2. problem occurs in several places in the program
3. this problem is widespread throughout the program

6. Most heuristic evaluations involve 4 to 5 experts. Once all experts have completed their evaluations, they may be brought together for a debriefing led by a moderator. The discussion of the usability problems may be videotaped for further analysis. If major differences appear in the problems found or the ratings given, the moderator may try to get the experts to resolve their differences and reach consensus. The experts may also be asked to suggest strategies for resolving the major usability problems found.

7. A heuristic evaluation report should then be compiled. Bar charts, tables, and other illustrations should be used to display the results. If feasible, screen captures should also be incorporated into the report to illustrate major problems as well as suggested enhancements.

8. The most important component of the heuristic report is a set of recommendations for improving the usability of the e-learning program. These should be as specific as possible to provide the designers with the information they need to eliminate the problems and improve the e-learning program.

Figure 1. Protocol for heuristic evaluation of e-learning programs.

1. **Visibility of system status**: The e-learning program keeps the learner informed about what is happening, through appropriate feedback within reasonable time.
Sample questions to ask yourself:
a. When modules and other components of the e-learning (e.g., streaming video) are downloading, is the status of the download communicated clearly?
b. Is the user provided with information that indicates that the e-learning program is operating correctly?

2. **Match between system and the real world**: The e-learning program's interface employs words, phrases and concepts familiar to the learner or appropriate to the content, as opposed to system-oriented terms. Wherever possible, the e-learning program utilizes real-world conventions that make information appear in a natural and logical order.
Sample questions to ask yourself:
a. Does the e-learning program's interactive design utilize metaphors that are familiar to the learner or related to the specific content of the program?
b. Is the interface "user friendly," given the content of the program and its target audience?

3. **Error recovery and exiting**: The e-learning program allows the learner to recover from input mistakes and provides a clearly marked "exit" to leave the program without requiring the user to go through an extended dialogue.
Sample questions to ask yourself:
a. Does the e-learning program distinguish between input errors and cognitive errors, allowing easy recovery from the former always, and from the latter when it is pedagogically appropriate?
b. Does the program allow the learner to leave whenever desired, but easily return to the closest logical point in the program?

4. **Consistency and standards**: When appropriate to the content and target audience, the e-learning program adheres to general software conventions and is consistent in its use of different words, situations, or actions.
Sample questions to ask yourself:
a. If appropriate to the content and target audience, does the e-learning product adhere to widely recognized standards for software interactions (e.g., going back in a Web browser)?
b. If the e-learning program does not utilize common software conventions for interactions, are the novel interactions appropriate for the content and target audience?
c. Does the program maintain an appropriate level of consistency in its design from one part of the program to another?

5. **Error prevention**: The e-learning program is designed to prevent common problems from occurring in the first place.
Sample questions to ask yourself:
a. Is the program designed so that the learner recognizes when he/she has made a mistake related to input rather than content?
b. Is the e-learning program designed to provide a second chance when unexpected input is received (e.g., "You typed "bat" in response to the question. Did you mean "tab?")?

6. **Navigation support**: The e-learning program makes objects, actions, and options visible so that the user does not have to remember information when navigating from one part of the program to another. Instructions for use of the program are always visible or easily retrievable.
Sample questions to ask yourself:
a. Does the interface of the e-learning program speak for itself so that extensive consultation of a manual or other documentation does not interfere with learning?
b. Does the e-learning program provide user-friendly hints and/or clear directions when the learner requests assistance?
c. Does the e-learning program include a map or table of contents that allows the learner to see what has been seen and not seen?

BEST COPY AVAILABLE

7. Aesthetics: Screen displays do not contain information that is irrelevant, and "bells and whistles" are not gratuitously added to the e-learning program.
Sample questions to ask yourself:
a. Are the font choices, colors, and sizes consistent with good screen design recommendations for e-learning programs?
b. Does the e-learning program utilize white space and other screen design conventions appropriately?

8. Help and documentation: The e-learning program provides help and documentation that is readily accessible to the user when necessary. The help provides specific concrete steps for the user to follow. All documentation is written clearly and succinctly.
Sample questions to ask yourself:
a. Is help provided that is screen or context specific?
b. Is help or documentation available from any logical part of the e-learning program?
c. Is help or documentation written clearly?

Figure 2. Usability and instructional design heuristics for evaluation of e-learning programs.

9. Interactivity: The e-learning program provides content-related interactions and tasks that support meaningful learning.
Sample questions to ask yourself:
a. Does the e-learning program provide meaningful interactions for the user, rather than simply presenting long sections of text?
b. Does the e-learning program engage the learner in content-specific tasks to complete and problems to solve that take advantage of the state-of-the-art of e-learning capabilities?

10. Message Design: The e-learning program presents information in accord with sound information-processing principles.
Sample questions to ask yourself:
a. Is the most important information on the screen placed in the areas most likely to attract attention?
b. Does the e-learning program follow good information presentation guidelines for organization and layout?

11. Learning Design: The interactions in the e-learning program have been designed in accord with sound principles of learning theory.
Sample questions to ask yourself:
a. Does the e-learning program follow an appropriate learning design to achieve its stated objectives?
b. Does the e-learning program engage learners in tasks that are closely aligned with the learning goals and objectives?

12. Media Integration: The inclusion of media in the e-learning program serves clear pedagogical and/or motivational purposes.
Sample questions to ask yourself:
a. Is media included that is obviously superfluous, i.e., lacking a strong connection to the objectives and design of the program?
b. Is the most appropriate media selected to match message design guidelines or to support instructional design principles?
c. If appropriate to the content, are various forms media included for remediation and/or enrichment?

13. Instructional Assessment: The e-learning program provides assessment opportunities that are aligned with the program objectives and content.
Sample questions to ask yourself:
a. If appropriate to the content, does the e-learning program provide opportunities for self-assessments that advance learner achievement?
b. If appropriate to the content, do assessments provide sufficient feedback to the learner to provide remedial directions?
c. Wherever appropriate, are higher order assessments (e.g., analysis, synthesis, and evaluation) provided rather than lower order assessments (e.g., recall and recognition)?

14. Resources: The e-learning program provides access to all the resources necessary to support effective learning.
Sample questions to ask yourself:
a. Does the e-learning program provide access to a range of resources (e.g., examples or real data archives) appropriate to the learning context?

b. If the e-learning program includes links to external World Wide Web or Intranet resources, are the links kept up-to-date?
c. Are resources provided in a manner that replicates as closely as possible their availability and use in the real world?

15. Feedback: The e-learning program provides feedback that is contextual and relevant to the problem or task in which the learner is engaged.
Sample questions to ask yourself:
a. Is the feedback given at any specific time tailored to the content being studied, problem being solved, or task being completed by the learner?
b. Does feedback provide the learner with information concerning his/her current level of achievement within the program?
c. Does the e-learning program provide learners with opportunities to access extended feedback from instructors, experts, peers, or others through e-mail or other Internet communications?

Figure 2. Usability and instructional design heuristics for evaluation of e-learning programs (continued).

All seven doctoral students and the cognitive psychology faculty member served as the experts for this heuristic evaluation and followed the protocol during a one-week period in October 2001. After an initial group review of the findings, another week was allowed to refine the individual reports that each expert compiled. Next, the instructional technology (IT) professor led a debriefing session during which copies of all the reports were made available to all experts. During this session, an acceptable level of consensus was reached concerning the major usability and instructional design problems, their severity, and their extensiveness. After the discussion, the IT professor synthesized the findings of the individual heuristic evaluations into a succinct report that was provided to the client.

Findings

The results of the heuristic evaluation of the GMP Basics courseware identified a number of important, but resolvable, problems with the e-learning program. For example, one of the heuristics violated in the initial version of the GMP program was Number 3, Error Recover and Exiting. In Figure 3, the screen capture on the left is from a timeline game within the GMP Basics course that did not allow the learner to exit easily, short of quitting the program altogether. The screen capture on the right illustrates the same game in the new version of the e-learning program after being redesigned based upon the results of the heuristic evaluation described in this paper as well as the findings of a field evaluation conducted by the first author in December 2001. In the revised version of the program, the learner can easily go back to the previous screen or exit this part of the program.

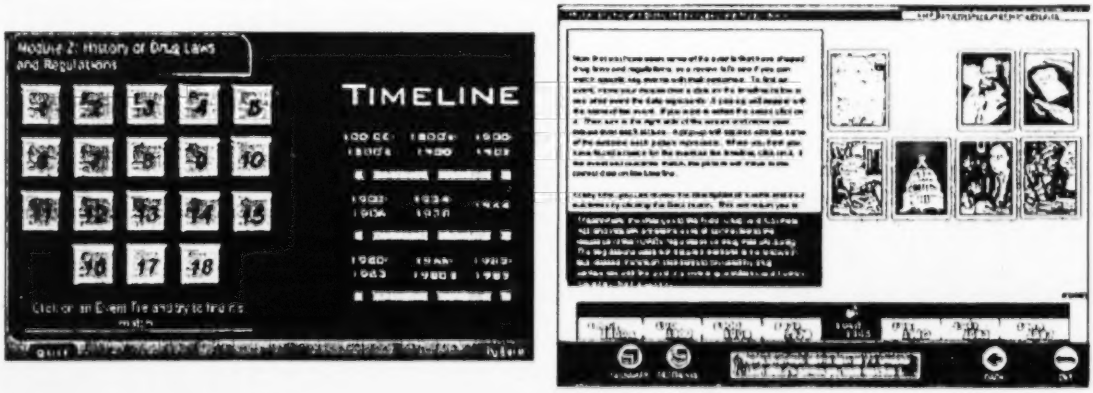

Figure 3. GMP Basics timeline game screen captures before (left) and after (right) evaluation.

Another heuristic violated in the early version of GMP Basics was Number 7, Aesthetics. In Figure 4, the screen capture on the left is from a quiz show game within first version of the GMP Basics course. The screen capture on the right illustrates the same game within the revised program. The aesthetics as well as the professional look-and-feel of the e-learning program have been enhanced in the new version.

BEST COPY AVAILABLE

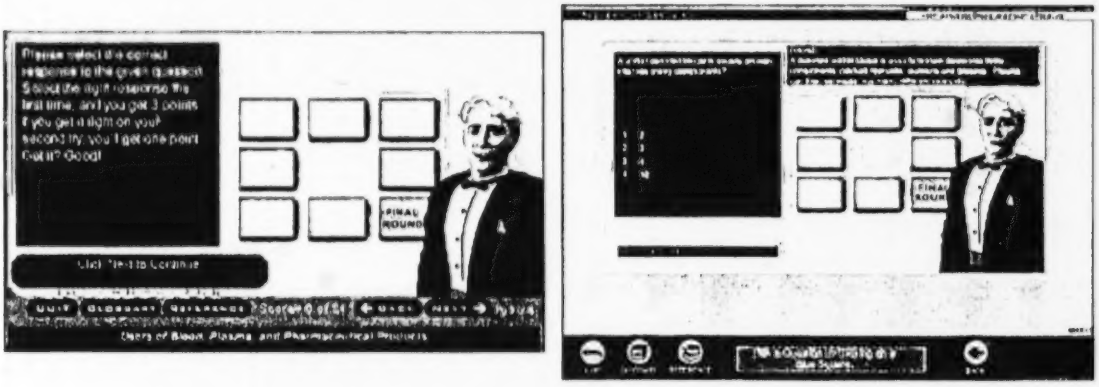

Figure 4. GMP Basics quiz show game screen captures before (left) and after (right) evaluation.

As a final illustration of the results of this heuristic evaluation, the Message Design (Heuristic Number 10) of the e-learning program was enhanced. In Figure 5, the screen capture on the left is from a typical instructional segment within the GMP Basics course that some learners found difficult to read because of the extensive use of light text on darker backgrounds. The screen capture on the right illustrates the same screen with the primary text appearing in a dark font on a lighter background. The screen on the right also illustrates improvements in the Navigation Support (Heuristic Number 6) that were made after this evaluation.

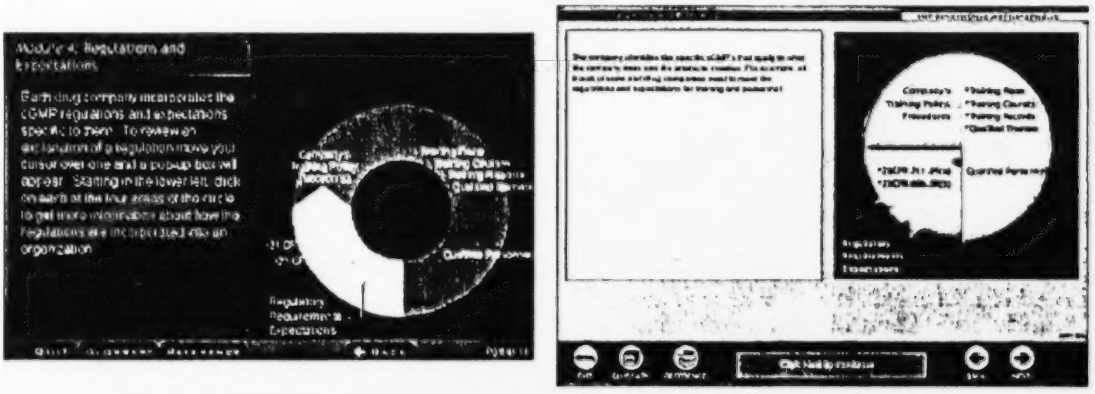

Figure 5. GMP Basics instructional screen captures before (left) and after (right) evaluation.

Interestingly, both the heuristic evaluation and the field evaluation revealed relatively few instructional design problems of the kind that would be represented by Heuristics Numbers 9 – 15. In other words, the original GMP program represented sound principles of instructional design, but the usability of that version of the program needed to be strengthened. Of course, usability is critical to a learner's experience with an e-learning program, and the effectiveness of even the most instructionally sound programs will be decreased if the learner's experience suffers from problems related to navigation, orientation, visual appeal, and other usability criteria (Reeves & Carter, 2001).

Discussion

There are limitations to this study that must be clarified. This heuristic evaluation took place over more than two weeks whereas most heuristic evaluations generally last only one or two days. There were eight experts involved instead of the usual four or five recommended by Nielsen (1994). In addition, changes made in the GMP Basics program were informed by a field evaluation conducted by the first author at a pharmaceutical manufacturing plant with real

BEST COPY AVAILABLE

workers, not by the heuristic evaluation alone. It is impossible to attribute exactly which enhancements were based upon which of the two evaluations. Nonetheless, LearnWright expressed great satisfaction with the results of the heuristic evaluation protocol and has requested that the protocol be applied to other programs being developed.

Although heuristic evaluation is fast, convenient, and economical, it is not sufficient for evaluating e-learning programs, and usability testing is advised (Nielsen, 2000). Heuristic evaluation and usability testing usually detect different types of usability problems (Nielsen, 1994). Heuristic evaluation is often conducted as a supplement to usability testing or when usability testing is not feasible. Through this study, we found that Nielsen's original ten usability heuristics were insufficient for e-learning programs, and we developed and refined the fifteen heuristics presented in Figure 2. We encourage others to adapt our e-learning heuristics to their own evaluation needs. We only request that, if possible, the results of subsequent evaluations be shared so that we can further refine this tool.

References

Nielsen, J. (2000). *Designing web usability: The practice of simplicity*. Indianapolis, IN: New Riders Publishing.

Nielsen, J. (1994). Heuristic evaluation. In Nielsen, J., and Mack, R.L. (Eds.), *Usability inspection methods*. New York: John Wiley & Sons.

Nielsen, J. (1993). *Usability engineering*. New York, NY: Academic Press, Inc.

Reeves, T. C., & Carter, B. J. (2001). Usability testing and return-on-investment studies: Key evaluation strategies for web-based training. In B. Khan (Ed.), *Web-based training* (pp. 547-557). Englewood Cliffs, NJ: Educational Technology Publications.

U.S. Department of Education
Office of Educational Research and Improvement (OERI)
National Library of Education (NLE)
Educational Resources Information Center (ERIC)

NOTICE

Reproduction Basis

X	This document is covered by a signed "Reproduction Release (Blanket)" form (on file within the ERIC system), encompassing all or classes of documents from its source organization and, therefore, does not require a "Specific Document" Release form.
	This document is Federally-funded, or carries its own permission to reproduce, or is otherwise in the public domain and, therefore, may be reproduced by ERIC without a signed Reproduction Release form (either "Specific Document" or "Blanket").